(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
19 September 2002 (19.09.2002)

PCT

(10) International Publication Number
WO 02/072756 A2

(51) International Patent Classification[7]: C12N

(21) International Application Number: PCT/US02/06818

(22) International Filing Date: 6 March 2002 (06.03.2002)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/274,013 7 March 2001 (07.03.2001) US

(63) Related by continuation (CON) or continuation-in-part (CIP) to earlier application:
US 60/274,013 (CON)
Filed on 7 March 2001 (07.03.2001)

(71) Applicant *(for all designated States except US)*: **MAYO FOUNDATION FOR MEDICAL EDUCATION AND RESEARCH** [US/US]; 200 First Street S.W., Rochester, MN 55905 (US).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **POESCHLA, Eric, M.** [US/US]; 1127 7th Street S.W., Rochester, MN 55902 (US). **WHITWAM, Todd, A.** [US/US]; P.O. Box 962, Rochester, MN 55903 (US).

(74) Agent: **ELLINGER, Mark, S.**; Fish & Richardson P.C., P.A., Sixty South Sixth Street, Suite 3300, Minneapolis, MN 55402 (US).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZM, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *without international search report and to be republished upon receipt of that report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 02/072756 A2

(54) Title: FELINE IMMUNODEFICIENCY VIRUS CENTRAL DNA FLAP

(57) Abstract: Method and materials relating to making and using feline immunodeficiency virus (FIV) lentiviral vectors are described, as well as host cells comprising such vectors.

FELINE IMMUNODEFICIENCY VIRUS CENTRAL DNA FLAP

BACKGROUND

1. Technical Field

The invention relates to methods and materials involved in making and using feline immunodeficiency virus (FIV) lentiviral vectors.

2. Background Information

Feline immunodeficiency virus (FIV) is a lentivirus that infects non-dividing cells, causes progressive CD4+ T cell depletion in domestic cats but not in humans, and has been used as a substrate for lentiviral vectors.

SUMMARY

The invention involves lentiviral vectors. Specifically, the invention provides for lentiviral vectors comprising a nucleotide sequence having at least 90% identity to the FIV central polypurine tract (cPPT) of any FIV strain or clone provided herein. Such lentiviral vectors can also include a central termination sequence (CTS) that is capable of functioning in conjunction with a cPPT to generate a central DNA flap required for nuclear import of a pre-integration complex through an intact nuclear membrane.

Unless otherwise defined, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this invention pertains. Although methods and materials similar or equivalent to those described herein can be used in the practice or testing of the present invention, suitable methods and materials are described below. All publications, patent applications, patents, and other references mentioned herein are incorporated by reference in their entirety. In case of conflict, the present specification, including definitions, will control. In addition, the materials, methods, and examples are illustrative only and not intended to be limiting.

Other features and advantages of the invention will be apparent from the following detailed description, and from the claims.

DESCRIPTION OF DRAWINGS

Figure 1 is a diagram depicting the location of the cPPT, the CTS, and the intervening sequence that comprise the central flap in FIV.

Figure 2A and 2B are bar graphs of infectious titers from either dividing (2A) or non-dividing (2B) cells. CT25a is the control vector lacking an FIV cPPT, CT26 is the vector containing an FIV cPPT, and CT26r is the vector containing an antisense FIV cPPT.

DETAILED DESCRIPTION

The invention provides methods and materials related to making and using lentiviral vectors. Specifically, the invention provides vectors comprising nucleic acids and host cells comprising lentiviral vectors or particles.

The experiments set out in the following Examples identify a central polypurine tract (cPPT) and a central termination sequence (CTS) that are involved in reverse transcription of FIV. The term "central polypurine tract" or "cPPT" as used herein refers to a nucleic acid sequence that comprises a plurality of contiguous purines (i.e., A or G) that can be interrupted by one or more pyrimidines (i.e., T, C, or U). For example, nucleotides 5-20 of the sequence set forth in SEQ ID NO: 1 is a cPPT. Typically, a cPPT is, without limitation, at least about 16 nucleotides in length. Further, a group of related cPPTs can contain a conserved pyrimidine. For example, the related FIV sequences set forth in SEQ ID NOs: 1, 2, 3, 4, 5, 6, and 7 contain a conserved thymidine located at position 4970 of the FIV genome (for FIV numbering conventions see Talbott *et al.*, *Proc. Nat. Acad. Sci. USA* 86:5743-7 (1989)), or, in other words, -2 to the G nucleotide that defines the 5′ terminus of the downstream (D+) strand.

The 5′ terminus of the D+ strand coincides with the 3′ boundary of a central gap located on the U+ strand (Figure 1). The term "central gap" as used herein refers to a single-stranded discontinuity located in the approximate center of an FIV genome. Such a central gap can be detected by practicing the methods provided herein, e.g., treating low molecular weight (LMW) DNA isolated from cells infected with FIV with S1 nuclease, an enzyme which cleaves single-stranded DNA selectively, and analyzing the resulting FIV genomic fragments with gel electrophoresis. Further, the term "gap nucleotide" as

used herein refers to a nucleotide that defines the 3′ boundary of a central gap, e.g., the 5′ terminus of the D+ strand. For example, the underlined "G" in the FIV sequence 5′-AAAAGAAGAGGTAGGA-3′ (nucleotides 5-20 of the sequence set forth in SEQ ID NO: 1) is the gap nucleotide that identifies the 5′ terminus of the D+ strand for the FIV strain comprising SEQ ID NO: 1.

The term "central termination sequence" or "CTS" as used herein refers to a site in a viral genome where upstream (U+) strand synthesis is terminated. For example, the FIV sequence 5′-CAAAAATT-3′ is a CTS (Figure 1).

The invention also provides intervening nucleic acid sequences. The term "intervening sequence" as used herein refers to a nucleic acid sequence flanked by a cPPT and a CTS. The cPPT and CTS can each be either 5′ or 3′ of the intervening sequence. For example, the cPPT can be 5′ of the intervening sequence, with the CTS being 3′ of the intervening sequence. An intervening sequence can vary in length, e.g., can be 8, 9, 10, 15, 20, 50, 100, 200, or 500 nucleotides in length.

The term "insert nucleic acid" as used herein refers to a nucleic acid sequence of interest that has been, or is to be, inserted into a transfer vector. An insert nucleic acid is distinct from a cPPT, a CTS, or an intervening sequence as they relate to the invention described herein. For example, a transfer vector can contain an insert nucleic acid encoding a marker polypeptide, e.g., green fluorescent protein (GFP), or any other polypeptide. Additionally, an insert nucleic acid can be a non-coding nucleic acid, e.g., an antisense nucleic acid.

The term "transfer vector" as used herein refers to a nucleic acid vector that has been adapted to receive an insert nucleic acid. A transfer vector within the scope of the invention includes vectors containing or not containing an insert nucleic acid.

Nucleic acids

The term "nucleic acid" as used herein encompasses RNA and DNA, including cDNA, genomic DNA, and synthetic (e.g., chemically synthesized) DNA. The nucleic acid can be double-stranded or single-stranded. Where single-stranded, the nucleic acid can be the sense strand or the antisense strand. In addition, nucleic acid can be circular or linear.

The term "isolated" as used herein with reference to nucleic acid refers to a naturally-occurring nucleic acid that is not immediately contiguous with both of the sequences with which it is immediately contiguous (one on the 5' end and one on the 3' end) in the naturally-occurring genome of the organism from which it is derived. For example, an isolated nucleic acid can be, without limitation, a recombinant DNA molecule of any length, provided one of the nucleic acid sequences normally found immediately flanking that recombinant DNA molecule in a naturally-occurring genome is removed or absent. Thus, an isolated nucleic acid includes, without limitation, a recombinant DNA that exists as a separate molecule (e.g., a cDNA or a genomic DNA fragment produced by PCR or restriction endonuclease treatment) independent of other sequences as well as recombinant DNA that is incorporated into a vector, an autonomously replicating plasmid, a virus (e.g., a retrovirus, adenovirus, or herpes virus), or into the genomic DNA of a prokaryote or eukaryote. In addition, an isolated nucleic acid can include a recombinant DNA molecule that is part of a hybrid or fusion nucleic acid sequence.

The term "isolated" as used herein with reference to nucleic acid also includes any non-naturally-occurring nucleic acid since non-naturally-occurring nucleic acid sequences are not found in nature and do not have immediately contiguous sequences in a naturally-occurring genome. For example, non-naturally-occurring nucleic acid such as an engineered nucleic acid is considered to be isolated nucleic acid. Engineered nucleic acid can be made using common molecular cloning or chemical nucleic acid synthesis techniques. Isolated non-naturally-occurring nucleic acid can be independent of other sequences, or incorporated into a vector, an autonomously replicating plasmid, a virus (e.g., a retrovirus, adenovirus, or herpes virus), or the genomic DNA of a prokaryote or eukaryote. In addition, a non-naturally-occurring nucleic acid can include a nucleic acid molecule that is part of a hybrid or fusion nucleic acid sequence.

It will be apparent to those of skill in the art that a nucleic acid existing among hundreds to millions of other nucleic acid molecules within, for example, cDNA or genomic libraries, or gel slices containing a genomic DNA restriction digest is not to be considered an isolated nucleic acid.

The term "exogenous" as used herein with reference to nucleic acid and a particular cell refers to any nucleic acid that does not originate from that particular cell as found in nature. Thus, all non-naturally-occurring nucleic acid is considered to be exogenous to a cell once introduced into the cell. It is important to note that non-naturally-occurring nucleic acid can contain nucleic acid sequences or fragments of nucleic acid sequences that are found in nature provided the nucleic acid as a whole does not exist in nature. For example, a nucleic acid molecule containing a genomic DNA sequence within an expression vector is non-naturally-occurring nucleic acid, and thus is exogenous to a cell once introduced into the cell, since that nucleic acid molecule as a whole (genomic DNA plus vector DNA) does not exist in nature. Thus, any vector, autonomously replicating plasmid, or virus (e.g., retrovirus, adenovirus, or herpes virus) that as a whole does not exist in nature is considered to be non-naturally-occurring nucleic acid. It follows that genomic DNA fragments produced by PCR or restriction endonuclease treatment as well as cDNAs are considered to be non-naturally-occurring nucleic acid since they exist as separate molecules not found in nature. It also follows that any nucleic acid containing a promoter sequence and polypeptide-encoding sequence (e.g., cDNA or genomic DNA) in an arrangement not found in nature is non-naturally-occurring nucleic acid.

Nucleic acid that is naturally-occurring can be exogenous to a particular cell. For example, an entire chromosome isolated from a cell of person X is an exogenous nucleic acid with respect to a cell of person Y once that chromosome is introduced into Y's cell.

The invention provides isolated nucleic acids that contain a nucleic acid sequence at least about 90% identical to nucleotides 5-20 of the sequence set forth in SEQ ID NO: 1, 2, 3, 4, 5, 6, or 7. Typically, the percent identity between two nucleic acid sequences is determined as follows. First, a nucleic acid sequence is compared to nucleotides 5-20 of a sequence set forth in SEQ ID NO: 1, 2, 3, 4, 5, 6, or 7 using the BLAST 2 Sequences (Bl2seq) program from the stand-alone version of BLASTZ containing BLASTN version 2.0.14 and BLASTP version 2.0.14. This stand-alone version of BLASTZ can be obtained from www.fr.com or www.ncbi.nlm.nih.gov. Instructions explaining how to use the Bl2seq program can be found in the readme file accompanying BLASTZ. Bl2seq performs a comparison between two sequences using either the BLASTN or BLASTP

algorithm. BLASTN is used to compare nucleic acid sequences, while BLASTP is used to compare amino acid sequences. To compare two nucleic acid sequences, the options are set as follows: -i is set to a file containing the first nucleic acid sequence to be compared (e.g., C:\seq1.txt); -j is set to a file containing the second nucleic acid sequence to be compared (e.g., C:\seq2.txt); -p is set to blastn; -o is set to any desired file name (e.g., C:\output.txt); -q is set to –1; -r is set to 2; and all other options are left at their default setting. For example, the following command can be used to generate an output file containing a comparison between two sequences: C:\Bl2seq –i c:\seq1.txt –j c:\seq2.txt –p blastn –o c:\output.txt –q –1 –r 2. If the two compared nucleic acid sequences share homology, then the designated output file will present those regions of homology as aligned sequences. If the two compared nucleic acid sequences do not share homology, then the designated output file will not present aligned sequences. Once aligned, the number of matches is determined by counting the number of positions where an identical nucleotide or amino acid residue is presented in both sequences.

The percent identity is determined by dividing the number of matches by the length of nucleotides 5-20 of the sequence set forth in SEQ ID NO: 1, 2, 3, 4, 5, 6, or 7, followed by multiplying the resulting value by 100. For example, if (1) a sequence is compared to nucleotides 5-20 of the sequence set forth in SEQ ID NO: 1, and (2) the number of matches is 8, then the sequence being compared to nucleotides 5-20 of the sequence set forth in SEQ ID NO: 1 has a percent identity of 50 (i.e., 8 ÷ 16 * 100 = 50) to nucleotides 5-20 of the sequence set forth in SEQ ID NO: 1.

It is noted that the percent identity value is rounded to the nearest tenth. For example, 78.11, 78.12, 78.13, and 78.14 is rounded down to 78.1, while 78.15, 78.16, 78.17, 78.18, and 78.19 is rounded up to 78.2. It is also noted that the length value will always be an integer.

The invention also provides isolated nucleic acid molecules that are at least about 5 bases in length (e.g., at least about 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 40, 50, 60, 100, 250, 500, 750, 1000, 1500, 2000, 3000, 4000, or 5000 bases in length) and hybridize, under hybridization conditions, to the sense or antisense strand of a nucleic acid having a sequence set forth in SEQ ID NO: 1, 2, 3, 4, 5, 6, or 7. The hybridization conditions can be moderately or highly stringent hybridization conditions.

For the purpose of this invention, moderately stringent hybridization conditions mean the hybridization is performed at about 42°C in a hybridization solution containing 25 mM KPO_4 (pH 7.4), 5X SSC, 5X Denhart's solution, 50 μg/mL denatured, sonicated salmon sperm DNA, 50% formamide, 10% Dextran sulfate, and 1-15 ng/mL probe (about $5x10^7$ cpm/μg), while the washes are performed at about 50°C with a wash solution containing 2X SSC and 0.1% sodium dodecyl sulfate.

Highly stringent hybridization conditions mean the hybridization is performed at about 42°C in a hybridization solution containing 25 mM KPO_4 (pH 7.4), 5X SSC, 5X Denhart's solution, 50 μg/mL denatured, sonicated salmon sperm DNA, 50% formamide, 10% Dextran sulfate, and 1-15 ng/mL probe (about $5x10^7$ cpm/μg), while the washes are performed at about 65°C with a wash solution containing 0.2X SSC and 0.1% sodium dodecyl sulfate.

Isolated nucleic acid molecules within the scope of the invention can be obtained using any method including, without limitation, common molecular cloning and chemical nucleic acid synthesis techniques. For example, PCR can be used to obtain an isolated nucleic acid molecule containing a nucleic acid sequence sharing similarity to nucleotides 5-20 of the sequence set forth in SEQ ID NO: 1. PCR refers to a procedure or technique in which target nucleic acid is amplified in a manner similar to that described in U.S. Patent No. 4,683,195, and subsequent modifications of the procedure described therein. Generally, sequence information from the ends of the region of interest or beyond are used to design oligonucleotide primers that are identical or similar in sequence to opposite strands of a potential template to be amplified. Using PCR, a nucleic acid sequence can be amplified from RNA or DNA. For example, a nucleic acid sequence can be isolated by PCR amplification from total cellular RNA, total genomic DNA, and cDNA as well as from bacteriophage sequences, plasmid sequences, viral sequences, and the like. When using RNA as a source of template, reverse transcriptase can be used to synthesize complimentary DNA strands.

Isolated nucleic acid molecules within the scope of the invention also can be obtained by mutagenesis. For example, an isolated nucleic acid containing nucleotides 5-20 of the sequence set forth in SEQ ID NO: 1 can be mutated using common molecular cloning techniques (e.g., site-directed mutagenesis). Possible mutations include, without

limitation, deletions, insertions, and substitutions, as well as combinations of deletions, insertions, and substitutions.

In addition, nucleic acid and amino acid databases (e.g., GenBank®) can be used to obtain an isolated nucleic acid molecule within the scope of the invention. For example, any amino acid sequence having some homology to a sequence set forth in SEQ ID NO: 8, 9, or 10 can be used as a query to search GenBank®.

Further, nucleic acid hybridization techniques can be used to obtain an isolated nucleic acid molecule within the scope of the invention. Briefly, any nucleic acid molecule having some homology to nucleotides 5-20 of the sequence set forth in SEQ ID NO: 1, 2, 3, 4, 5, 6, or 7 can be used as a probe to identify a similar nucleic acid by hybridization under conditions of moderate to high stringency. Once identified, the nucleic acid molecule then can be purified, sequenced, and analyzed to determine whether it is within the scope of the invention as described herein.

Hybridization can be done by Southern or Northern analysis to identify a DNA or RNA sequence, respectively, which hybridizes to a probe. The probe can be labeled with a biotin, digoxygenin, an enzyme, or a radioisotope such as ^{32}P. The DNA or RNA to be analyzed can be electrophoretically separated on an agarose or polyacrylamide gel, transferred to nitrocellulose, nylon, or other suitable membrane, and hybridized with the probe using standard techniques well known in the art such as those described in sections 7.39-7.52 of Sambrook *et al.*, (1989) Molecular Cloning, second edition, Cold Spring harbor Laboratory, Plainview, NY. Typically, a probe is at least about 16 nucleotides in length. For example, a probe corresponding to nucleotides 5-20 of the sequence set forth in SEQ ID NO: 1 can be used to identify an identical or similar nucleic acid. In addition, probes longer or shorter than 16 nucleotides can be used.

Host cells

A host cell within the scope of the invention is any cell containing at least one isolated nucleic acid molecule described herein. Such cells can be prokaryotic and eukaryotic cells. It is noted that cells containing an isolated nucleic acid molecule within the scope of the invention are not required to express a polypeptide. In addition, the isolated nucleic acid molecule can be integrated into the genome of the cell or maintained

in an episomal state. Thus, host cells can be stably or transiently transfected with a construct containing an isolated nucleic acid molecule of the invention.

Host cells within the scope of the invention can contain an exogenous vector that comprises a nucleotide sequence having 90% identity to nucleotides 5-20 of the sequence set forth in SEQ ID NO: 1, 2, 3, 4, 5, 6, or 7. For example, a somatic cell containing a vector comprising nucleotides 5-20 of the sequence set forth in SEQ ID NO: 1 is a host cell. Additionally, cultured cells, e.g., Chinese hamster ovary (CHO) cells, COS-1 cells, human embryonic kidney 293 cells, NIH3T3 cells, BHK-21 cells, MDCK cells, or human vascular endothelial cells (HUVEC), containing an exogenous vector that comprises a nucleotide sequence having 90% identity to nucleotides 5-20 of the sequence set forth in SEQ ID NO: 1, 2, 3, 4, 5, 6, or 7 can be host cells.

Any methods can be used to introduce an isolated nucleic acid molecule into a cell *in vivo* or *in vitro*. For example, calcium phosphate precipitation, electroporation, heat shock, lipofection, microinjection, and viral-mediated nucleic acid transfer are common methods that can be used to introduce an isolated nucleic acid molecule into a cell. In addition, naked DNA can be delivered directly to cells *in vivo* as describe elsewhere (U.S. Patent Number 5,580,859 and U.S. Patent Number 5,589,466 including continuations thereof). Further, isolated nucleic acid molecules can be introduced into cells by generating transgenic animals.

Transgenic animals can be aquatic animals (such as fish, sharks, dolphin, and the like), farm animals (such as pigs, goats, sheep, cows, horses, rabbits, and the like), rodents (such as rats, guinea pigs, and mice), non-human primates (such as baboon, monkeys, and chimpanzees), and domestic animals (such as dogs and cats). Several techniques known in the art can be used to introduce isolated nucleic acid molecules into animals to produce the founder lines of transgenic animals. Such techniques include, without limitation, pronuclear microinjection (U.S. Patent No. 4,873,191); retrovirus mediated gene transfer into germ lines (Van der Putten *et al.*, *Proc. Natl. Acad. Sci., USA*, 82:6148 (1985)); gene transfection into embryonic stem cells (Gossler A *et al.*, *Proc Natl Acad Sci USA* 83:9065-9069 (1986)); gene targeting into embryonic stem cells (Thompson *et al.*, *Cell*, 56:313 (1989)); nuclear transfer of somatic nuclei (Schnieke AE *et al.*, *Science* 278:2130-2133 (1997)); and electroporation of embryos (Lo CW, *Mol.*

Cell. Biol., 3:1803-1814 (1983)). Once obtained, transgenic animals can be replicated using traditional breeding or animal cloning.

Any method can be used to identify cells containing an isolated nucleic acid molecule of the invention. Such methods include, without limitation, PCR and nucleic acid hybridization techniques such as Northern and Southern analysis. In some cases, immunohistochemistry and biochemical techniques can be used to determine if a cell contains a particular isolated nucleic acid molecule by detecting the expression of a polypeptide encoded by that particular nucleic acid molecule.

The invention will be further described in the following examples, which do not limit the scope of the invention described in the claims.

EXAMPLES

Example 1 – Recovery of FIV pre-integration complexes by single-round infection

A single round infection assay that employs a VSV-G pseudotyped, *env* frame-shifted modification of CT5 (CT5efs) was used to maximize both the yields of produced virus and the peak yield of FIV pre-integration complexes in target cells. For a description of CT5 see Poeschla and Looney *J. Virol.* 72:6858-6866 (1998). CT5efs contains a 29 nucleotide insertion in the *env* ORF, resulting in a frame shift without deletion of any FIV sequences. This construct, therefore, permitted unbiased analysis in the present study because it abrogates Env-induced cytopathicity but does not delete any potential internal plus strand initiation sites.

VSV-G pseudotyped, replication-defective CT5efs particles (CT5efs(VSV-G)) produced in 293T cells were titered on Crandell feline kidney cells (CrFK) cells using immunoperoxidase staining for Gag/Pol. The end-point dilution titer of unconcentrated CT5efs(VSV-G) on CrFK cells was found to be 0.9×10^7/ ml. To maximize pre-integration complex generation, uninfected CrFK cells were plated at a 3:1 ratio with chronically infected CrFK cells. Six hours later the replication-defective CT5efs(VSV-G) was used to infect a culture of target cells at an m.o.i of 7.5. No cytopathicity was observed in the target cells prior to the harvesting of DNA samples.

FIV pre-integration complexes were detected by Southern blotting as a provirus-sized band in Hirt extracts of these cells. These data demonstrate that the described

single-round infection assay can be used to recover FIV pre-integration complexes in diagnostic amounts from infected target cells.

Example 2 - Detection of a plus strand discontinuity

Low molecular weight (LMW) DNA was isolated by Hirt extraction from the cultures described in Example 1 at 12, 24, 42, 48 and 60 hours after infection. The resulting isolated LMW DNA contained reverse-transcribed genomes lacking any integrated viral DNA. Simultaneous control Hirt extracts were made from uninfected CrFK cells. The extracted LMW DNA samples were digested first with Spe I (40 U/mg DNA) for one hour at 37°C in buffer containing 250 mM M NaCl, 50 mM $C_2H_3O_2Na$, 1mM $ZnSO_4$, 50mg/ml BSA. Following Spe I digestion, the LMW DNA samples were further treated with S1 nuclease (4U/mg of DNA) for one hour at 37°C in buffer containing 250 mM M NaCl, 50 mM $C_2H_3O_2Na$, 1mM $ZnSO_4$, 50mg/ml BSA. The resulting digested LMW DNA samples were separated on a 1.2% agarose gel by electrophoresis and analyzed by Southern blotting. The probe was a ^{32}P labeled Bgl II-Spe I restriction fragment spanning nucleotide 6457 to 8287 of FIV *env*. For a description of the probe see Talbott *et al.*, *Proc. Nat. Acad. Sci. USA* 86:5743-7 (1989).

An 8-10 kb band representing reverse-transcribed unintegrated proviral DNA was detected in the Hirt extracts from infected CrFK cells but not from uninfected CrFK cells. In addition, a band approximately 3.5 kb in size was detectable in the S1 nuclease digested LMW DNA from infected cells, but not in LMW DNA from uninfected cells or from infected cells in the absence of S1 nuclease treatment. The size of this 3.5kb band suggests a single strand gap in the center of the genome. The approximate location was estimated to be the 3' region of *pol*, a region consistent with the location of previously described cPPTs. These results demonstrate that the FIV reverse-transcribed genome contains a centrally-located single strand gap.

Example 3 – Primer extention analysis mapping the FIV central polypurine tract

LMW DNA was extracted from infected and uninfected cells as described in Example 2. 5' to 3' primer extension was performed using 10μg LMW DNA, 25 mM of each dNTP, 10 U Taq polymerase, 10mM Tris-HCl, 15mM $MgCl_2$, 50 mM KCl, and 10

pmol of a 5' end-labeled primer predicted to anneal to the plus strand about 100 bp 3' of the approximate location of the FIV central polypurine tract (cPPT). The sequence of the 5' end-labeled primer is 5'-ATAATAAATCCACTGTGC-3' (SEQ ID NO: 11). The primer extention reaction conditions were as follows: denaturing at 95°C for 30 seconds, annealing at 45°C for 30 seconds, and elongating at 72° for 60 seconds. The reactions were cycled 30 times. A 5 minute denaturation step at 95°C preceded the first reaction cycle. Following primer extention, the reactions were stopped with one-half reaction volume of formamide loading dye. The primer extension reactions were separated by electrophoresis on a 6% acrylamide / 7M urea gel in parallel with Sanger sequencing reactions generated from the CT5 plasmid using the same end-labeled primer (SEQ ID NO: 11).

A single discrete stop in primer extension was seen in LMW DNA samples from infected cells but not in LMW DNA samples from uninfected control cells. This stop maps the 3' boundary of a putative gap (i.e., the 5' terminus of the internally initiated D+ strand) to a G residue in a purine-rich tract in *pol* (nucleotide position 4972; underlined G in: 5'-AAAAGAAGAGGTAGGA-3'; nucleotides 5-20 of the sequence set forth in SEQ ID NO: 1). This identified 16 nucleotide cPPT is located centrally within the FIV genome (nucleotide positions 4959-4974). The D+ strand origin at nucleotide position 4972 is 235 nucleotides 3' of the precise center of the 9474 nucleotide provirus and 269 nucleotides 5' of the terminus of *pol*. The cPPT is a run of mostly purines in the distal portion of the integrase gene. A comparison of this sequence with the same region in other FIV strains is shown in Table 1. An invariant pyrimidine (a thymidine) is located –2 to the gap nucleotide, and some FIV cPPTs contain an additional pyrimidine (a C nucleotide). This thymidine (uracil in the RNA primer) is neither required nor optimal for the glycine codon (GGU) at this position. As in other lentiviral genomes, codon usage for the 88 glycines in FIV *pol* is strongly biased toward a purine, particularly adenine, at the third position (GGU: 15%; GGA: 64%; GGC: 5%; GGG: 17%). GGU is also the least favored glycine codon in overall mammalian usage (GGU: 16%; GGA: 24%; GGC: 36%; GGG: 24%). These data demonstrate that the FIV genome contains a cPPT. These data further suggest that cPPTs from various FIV strains share some common features.

Table 1: Alignment of the Identified FIV 34TF10 cPPT With Polypurine Tract Regions in other FIV strains

FIV cPPT Comparisons			
Virus/ clone	**Function mapped***	**Integrase alignment (cPPT sequence in bold-face)**	**Predicted AA sequence*** **
FIV34Tf10	Yes**	[5]**ttttaaaagaagaggtaggatagga** (SEQ ID NO: 1)	**KRRGRIG** (SEQ ID NO: 8)
FIVPPR	No	**ttttaaacaaaggggtagaatagga** (SEQ ID NO: 2)	**KQRGRIG** (SEQ ID NO: 9)
FIVNCSUJSY	No	**ctttaaacaaaggggtagactagga** (SEQ ID NO: 3)	**KQRGRLG** (SEQ ID NO: 10)
FIV Oma	No	**ctttaaacaaaggggtagaataggg** (SEQ ID NO: 4)	**KQRGRIG** (SEQ ID NO: 9)
FIVZ1	No	**ttttaaaagaagaggtaggatagga** (SEQ ID NO: 5)	**KRRGRIG** (SEQ ID NO: 8)
FIVTM2	No	**ttttaaacaaaggggtagactaggg** (SEQ ID NO: 6)	**KQRGRLG** (SEQ ID NO: 10)
FIVGVEPX	No	**ttttaaacaaaggggtagactaggg** (SEQ ID NO: 7)	**KQRGRLG** (SEQ ID NO: 10)

*Denotes whether any cPPT function in reverse transcription has been identified for the listed sequence; references given in parentheses. **Mapped in the present study. ***Predicted amino acid sequence encoded by bold-faced nucleotides. [5]Numbering of the nucleic acid sequences in Table 1 starts with the left-most nucleotide as nucleotide number 1. For example, the bold-faced nucleotides in SEQ ID NOs :1-7 are nucleotides 5-20.

Example 4 - RACE PCR analysis mapping the FIV central termination sequence

1 µg of LMW DNA samples extracted from infected cells and from uninfected control cells as described in Example 2 was denatured at 95°C for 4 minutes and cooled on ice. The LMW DNA samples were then poly(dA) tailed by terminal deoxynucleotide transferase. The tailing reactions contained 50 units terminal deoxynucleotide transferase, 5 µM dATP, 0.75 mM cobalt chloride, 200 mM potassium cacodylate, 0.25 mg/ml BSA, and 25 mM Tris HCl, pH 6.6, in a 20 µl reaction volume. The tailing reactions were incubated at 37°C for 30 minutes. The resulting poly(dA) tailed products were extracted with phenol:chloroform:isoamyl alcohol and then ethanol precipitated with glycogen as carrier. 6.67 µL of each precipitated DNA product was used as template in a PCR using a sense primer FIV4576 (5'-CTGGTATCTGGCAAATGGATTGC-3'; SEQ ID NO: 12) and an antisense primer with a 17 bp oligo(dT) sequence (5'-TCTAGACCATGGAGATCTCGATCGTTTTTTTTTTTTTTTT-3'; SEQ ID NO: 13)

for 25 cycles (denaturing at 95°C for 30 seconds, annealing at 56°C for 30 seconds, and elongating at 72°C for 90 seconds). The PCR mixtures contained 1.5 mM $MgCl_2$, 0.2 mM dNTPs, and Platinum Taq (Gibco BRL). Following PCR, 0.1 % of the resulting PCR product was used as template in a nested PCR using either FIV4702 (5'-CTGTCTTACAATTGTTGAGTGC-3'; SEQ ID NO: 14) or FIV4674 (5'-CAAGAAACTGCTGACTGTACAG-3'; SEQ ID NO: 15) as the sense primer and the same oligo(dT) antisense primer (5'-TCTAGACCATGGAGATCTCGATCGTTTTTTTTTTTTTTTT-3'; SEQ ID NO: 13) for 30 cycles (denaturing at 95°C for 30 seconds, annealing at 56°C for 30 seconds, and elongating at 72°C for 90 seconds). Following the nested PCR, the resulting products were directly cloned without any fractionation or purification into the pCR2.1 vector using the TA cloning method (Invitrogen). The cloned inserts were then sequenced on an ABI Prism 377 automated sequencer to ensure that no spurious mutations had been introduced during PCR. Following insert sequencing, the vectors containing cloned inserts were introduced into DH5α bacteria by transformation. Transformed bacteria were then plated onto agar containing ampicillin such that individual ampicillin-resistant clones could be isolated 24 hours after.

A total of 24 randomly picked insert-containing clones from two separate terminal transferase tailings and nested PCR amplification experiments with different nesting primers were sequenced. Prior to cloning, PCR products were composed of single prominent bands that placed the site of joining of the oligod(dA) tail approximately 80-120 nucleotides downstream of the cPPT depending on the exact site of oligo(dT) annealing to the tail. A very faint band about 200 nucleotides longer than the main band, possibly representing a minor U+ strand termination, can be seen in both panels but was not detected in any sequenced clones. No products were detected in amplifications of LMW DNA from uninfected cells. The precise 3' terminus of the U+ strand was determined in all 24 clones to occur at the second T nucleotide in a CA_5T_2 sequence downstream of the cPPT. These data demonstrate that the FIV genome contains a central termination sequence with a sequence of 5'-CAAAAATT-3'.

Example 5 – The FIV cPPT increases the efficiency of FIV vectors

A 301 bp FIV region extending from 66 nucleotides upstream of the plus strand gap to 234 nucleotides downstream of the gap was amplified by PCR. The resulting amplification product was inserted into an FIV-based lentiviral vector that contained no *pol* sequences and only the first 311 nucleotides of the *gag* ORF. Stocks of the parental CT25a vector lacking the FIV cPPT were simultaneously prepared, and all vector stocks were equilibrated for reverse transcriptase activity. The endpoint dilution titers of cPPT-containing vectors were approximately 4.5 fold greater than the titers for the cPPT-minus vector. These data demonstrate that the FIV cPPT increases the efficiency of FIV vectors.

OTHER EMBODIMENTS

It is to be understood that while the invention has been described in conjunction with the detailed description thereof, the foregoing description is intended to illustrate and not limit the scope of the invention, which is defined by the scope of the appended claims. Other aspects, advantages, and modifications are within the scope of the following claims.

WHAT IS CLAIMED IS:

1. A vector comprising a nucleotide sequence having at least 90% identity to nucleotides 5-20 of the sequence set forth in SEQ ID NO: 1, 2, 3, 4, 5, 6, or 7, said vector being free of nucleotide sequence encoding a functional viral integrase.

2. The vector of claim 1, further comprising a central termination sequence and an intervening sequence.

3. An FIV vector comprising an intervening sequence flanked by a central polypurine tract and a central termination sequence, wherein said FIV vector is free of nucleotide sequence encoding a functional viral integrase.

4. The FIV vector of claim 3, wherein said intervening sequence is at least about 8 nucleotides to at least about 100 nucleotides in length.

5. The FIV vector of claim 3, wherein said intervening sequence is at least 88 nucleotides in length.

6. The FIV vector of claim 3, wherein said central polypurine tract is at least 90% identical to nucleotides 5-20 of the sequence set forth in SEQ ID NO: 1, 2, 3, 4, 5, 6, or 7.

7. The FIV vector of claim 3, wherein said central polypurine tract comprises nucleotides 5-20 of the sequence set forth in SEQ ID NO: 1, 2, 3, 4, 5, 6, or 7

8. The FIV vector of claim 3, wherein said central polypurine tract is 5'-AAAAGAAGAGGTAGGA-3' (nucleotides 5-20 of the sequence set forth in SEQ ID NO: 1).

9. The FIV vector of claim 3, wherein said central termination sequence is 5'-CAAAAATT-3'.

10. An FIV vector comprising a central polypurine tract encoding the amino acid sequence selected from the group consisting of KRRGRIG (SEQ ID NO: 8), KQRGRIG (SEQ ID NO: 9), and KQRGRLG (SEQ ID NO: 10).

11. An isolated nucleic acid comprising a retroviral nucleic acid sequence free of nucleic acid sequence encoding a functional viral integrase, wherein said retroviral nucleic acid sequence comprises a pyrimidine within a tract of otherwise contiguous purines, said tract being from at least about 8 purines to at least about 20 purines in length, said pyrimidine located at either position –2 or –3 relative to a gap nucleotide in said tract.

12. An FIV transfer vector comprising an intervening sequence, a central polypurine tract, and a central termination sequence, said intervening sequence being flanked by said central polypurine tract and said central termination sequence, wherein said FIV transfer vector is free of nucleotide sequence encoding a functional viral integrase.

13. The transfer vector of claim 12, wherein said transfer vector further comprises an insert nucleic acid.

14. The transfer vector of claim 12, wherein said intervening sequence is at least about 8 nucleotides to at least about 100 nucleotides in length.

15. The transfer vector of claim 12, wherein said intervening sequence is at least 88 nucleotides in length.

16. The transfer vector of claim 12, wherein said central polypurine tract is at least 90% identical to nucleotides 5-20 of the sequence set forth in SEQ ID NO: 1, 2, 3, 4, 5, 6, or 7.

17. The transfer vector of claim 12, wherein said central polypurine tract comprises nucleotides 5-20 of the sequence set forth in SEQ ID NO: 1, 2, 3, 4, 5, 6, or 7.

18. The transfer vector of claim 12, wherein said central polypurine tract is 5′-AAAAGAAGAGGTAGGA-3′ (nucleotides 5-20 of the sequence set forth in SEQ ID NO: 1)

19. The transfer vector of claim 12, wherein said central termination sequence is 5′-CAAAAATT-3′.

20. A host cell comprising a vector, said vector comprising a nucleotide sequence having at least 90% identity to nucleotides 5-20 of the sequence set forth in SEQ ID NO: 1, 2, 3, 4, 5, 6, or 7, said vector being free of nucleotide sequence encoding a functional viral integrase.

21. A host cell comprising an FIV vector, said FIV vector comprising an intervening sequence flanked by a central polypurine tract and a central termination sequence, wherein said FIV vector is free of nucleotide sequence encoding a functional viral integrase.

22. A host cell comprising an FIV vector, said FIV vector comprising a central polypurine tract encoding the amino acid sequence selected from the group consisting of KRRGRIG (SEQ ID NO: 8), KQRGRIG (SEQ ID NO: 9), and KQRGRLG (SEQ ID NO: 10).

23. A host cell comprising an isolated nucleic acid, said isolated nucleic acid comprising a retroviral nucleic acid sequence free of nucleic acid sequence encoding a functional viral integrase, wherein said retroviral nucleic acid sequence comprises a pyrimidine within a tract of otherwise contiguous purines, said tract being from at least about 8 purines to at least about 20 purines in length, said pyrimidine located at either position –2 or –3 relative to a gap nucleotide in said tract.

24. A host cell comprising an FIV transfer vector, said transfer vector comprising an intervening sequence, a central polypurine tract, and a central termination sequence, said intervening sequence being flanked by said central polypurine tract and said central

termination sequence, wherein said FIV transfer vector is free of nucleotide sequence encoding a functional viral integrase.

25. The host cell of claim 24, wherein said transfer vector further comprises an insert nucleic acid.

26. A vector particle comprising FIV reverse transcriptase and a nucleic acid vector, said nucleic acid vector comprising an intervening sequence, a central polypurine tract, and a central termination sequence, said intervening sequence being flanked by said central polypurine tract and said central termination sequence, wherein said vector particle is free of nucleotide sequence encoding a functional viral integrase.

27. The vector particle of claim 26, wherein said nucleic acid vector further comprises an insert nucleic acid.

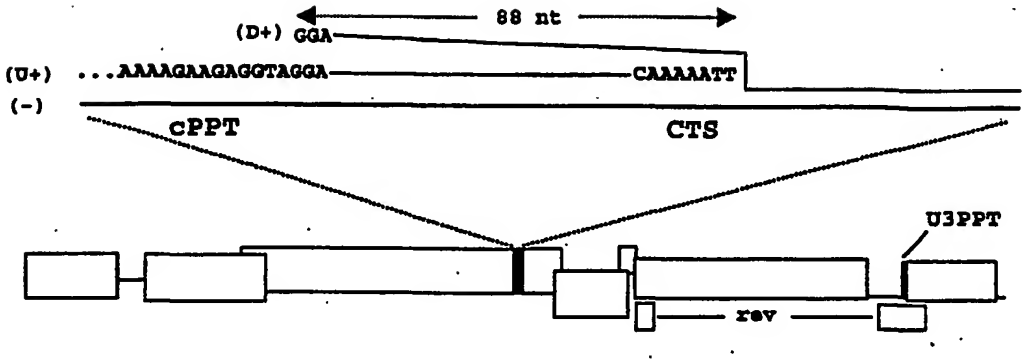

Figure 1

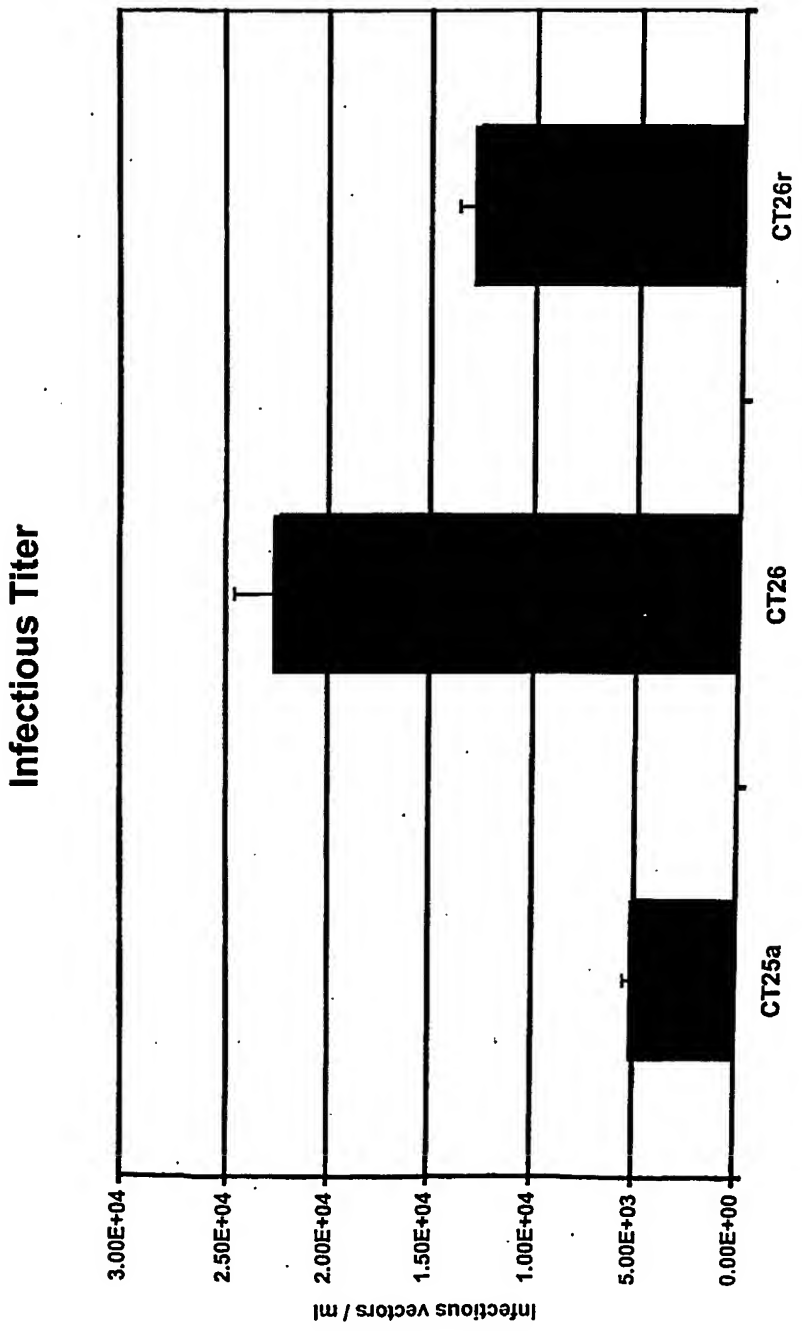
Infectious Titer
Infectious vectors / ml
3.00E+04
2.50E+04
2.00E+04
1.50E+04
1.00E+04
5.00E+03
0.00E+00
CT26r
CT26
CT25a

Figure 2A

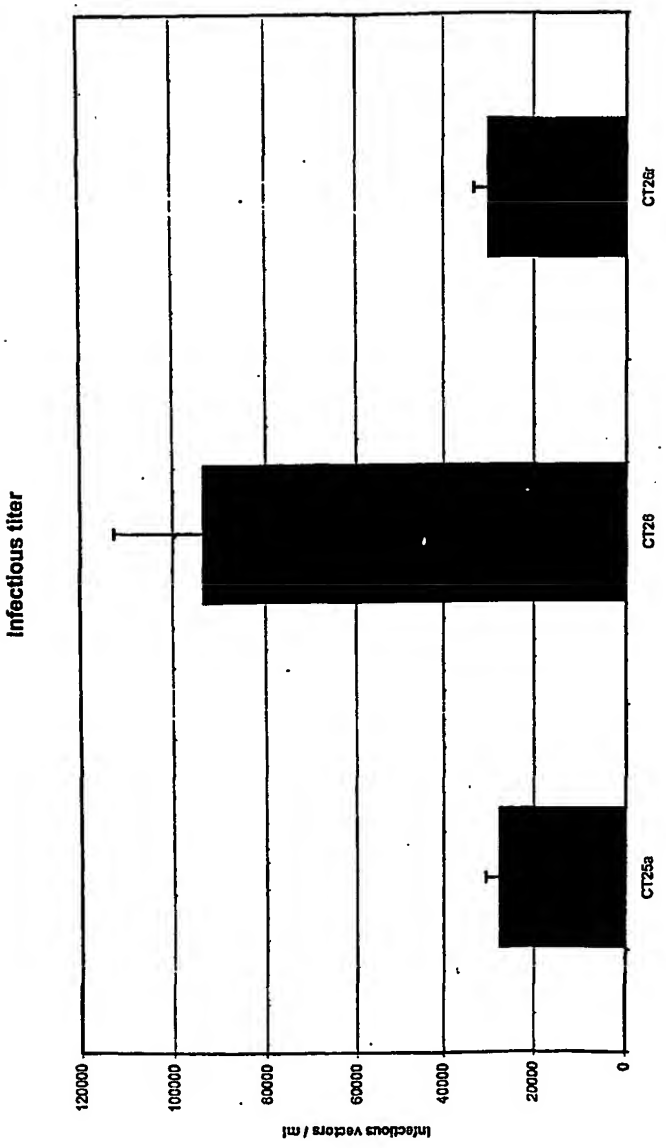
Infectious titer
Infectious vectors / ml
120000
100000
80000
60000
40000
20000
0
CT25a
CT26
CT26r

Figure 2B